Mitigating the Limited View Problem in Photoacoustic Tomography for a Planar Detection Geometry by Regularized Iterative Reconstruction

Jiaqi Zhu, Nam Huynh, Olumide Ogunlade, Rehman Ansari, Felix Lucka, Ben Cox, and Paul Beard

***Abstract*— The use of a planar detection geometry in photoacoustic tomography results in the so-called limited-view problem due to the finite extent of the acoustic detection aperture. When images are reconstructed using one-step reconstruction algorithms, image quality is compromised by the presence of streaking artefacts, reduced contrast, image distortion and reduced signal-to-noise ratio. To mitigate this, model-based iterative reconstruction approaches based on least squares minimisation with and without total variation regularization were evaluated using *in-silico*, experimental phantom, *ex vivo* and *in vivo* data. Compared to one-step reconstruction methods, it has been shown that iterative methods provide better image quality in terms of enhanced signal-to-artefact ratio, signal-to-noise ratio, amplitude accuracy and spatial fidelity. For the total variation approaches, the impact of the regularization parameter on image feature scale and amplitude distribution was evaluated. In addition, the extent to which the use of Bregman iterations can compensate for the systematic amplitude bias introduced by total variation was studied. This investigation is expected to inform the practical application of model-based iterative image reconstruction approaches for improving photoacoustic image quality when using finite aperture planar detection geometries.**

***Index Terms*— Photoacoustic image reconstruction, planar detection geometry, iterative image reconstruction, total variation regularization, Bregman iteration.**

Manuscript received 30 January 2023; revised 9 March 2023; accepted 21 March 2023. Date of publication 28 April 2023; date of current version 31 August 2023. This work was supported in part by the European Research Council under Grant 741149; in part by the Cancer Research, U.K., in part by the Engineering and Physical Sciences Research Council; in part by the National Institute for Health Research, University College London Hospital; and in part by the Biomedical Research Centre. *(Corresponding author: Jiaqi Zhu.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of University College London under Application No. 1133/001.

Jiaqi Zhu, Rehman Ansari, and Ben Cox are with the Department of Medical Physics and Biomedical Engineering, University College London, WC1E 6BT London, U.K. (e-mail: rmapzhu@ucl.ac.uk; r.abdul@ucl.ac.uk; b.cox@ucl.ac.uk).

Nam Huynh, Olumide Ogunlade, and Paul Beard are with the Department of Medical Physics and Biomedical Engineering and the Wellcome EPSRC Centre for Interventional and Surgical Sciences, University College London, London WC1E 6BT, U.K. (e-mail: n.huynh@ucl.ac.uk; o.ogunlade@ucl.ac.uk; paul.beard@ucl.ac.uk).

Felix Lucka is with Centrum Wiskunde & Informatica, 1098 XG Amsterdam, The Netherlands (e-mail: Felix.Lucka@cwi.nl).

Digital Object Identifier 10.1109/TMI.2023.3271390

I. Introduction

PHOTOACOUSTIC imaging which utilises both optical and ultrasound energy is an emerging modality [1], [2], [3] with a broad range of potential pre-clinical [4], [5], [6], [7], [8], [9], [10] and clinical applications [11], [12], [13], [14], [15]. In photoacoustic tomography, widefield pulsed laser light is delivered through the surface of tissue and absorbed by light-absorbing chromophores. The optical absorption produces a temperature rise and a corresponding pressure increase resulting in the generation of acoustic waves. The induced acoustic waves then propagate to the surface of the tissue and are detected at different spatial points with point-like detectors. The image reconstructed from the detected acoustic waves provides a representation of the distribution of absorbed optical energy density.

Theoretically, an exact photoacoustic image can be reconstructed when the detector array fully encloses the object. If the object is partially enclosed and there exists a so-called 'visible' region [16] where the normal ray from any point on the boundary of an absorbing feature intersects the detector array, it is also possible in principle to reconstruct an exact image in that region. However, many practical *in vivo* applications do not permit enclosure of the target tissue or organ and require the use of a planar detection array of limited aperture. In this case, the photoacoustic signals can only be recorded over a limited solid angular aperture and there is no visible region so an exact image can never be found. In this "limited view" scenario, image quality is compromised by artefacts [17], blurring, structural distortion and amplitude scaling errors. The extent of this image degradation depends on the image reconstruction algorithm used. Commonly used one-step (i.e. non-iterative) reconstruction algorithms for a planar detection geometry such as those based on k-space [18] or time reversal methods [19], [20], [21] can only give the exact solution if the detection surface has an infinite detection aperture. When data from a finite aperture is used to reconstruct the image, image quality is compromised. Model-based iterative methods form an image by iteratively adjusting the reconstructed image until it best matches the measured photoacoustic times series. They can be regarded as finding the image that best "explains" the data acquired over a finite aperture. In doing so, they implicitly account for the limited

TABLE I
ABBREVIATIONS OF RECONSTRUCTION METHODS

Abbreviation	Meaning
TR	Time reversal
TR+	Non-negativity constrained TR
TR+TVdenoising	TR post-processed by total variation denoising
iLS+	Non-negativity constrained Iterative least square
TV+	Non-negativity constrained total variation regularisation
TV+Breg	Bregman iteration enhanced TV+

view in a way that single-step methods do not and offer the prospect of improved image accuracy.

Model-based iterative image reconstruction methods have been investigated for removing artefacts and improving image quality of photoacoustic tomography for semi-circular [22], hemispherical [22], cylindrical [22], [23], arc-shape [24], [25] and linear [26], [27] detection geometries. For a planar detection geometry, iterative methods have been used to reconstruct images using spatially sub-sampled data [28], [29]. However, for fully sampled data, there remains a need for a comprehensive investigation of iterative reconstruction approaches and a rigorous assessment of their performance in terms of artefact level, geometrical distortion, amplitude accuracy and robustness to noise. In addition, an understanding of the impact of regularization parameters on image fidelity and signal-to-noise ratio (SNR) as a function of the target geometry and instrument noise is required. The aim of this study is to explore these factors and advance the practical application of model-based iterative reconstruction methods for improving the quality of images acquired using a planar detection geometry. This involved evaluating the efficacy and limitations of model-based iterative image reconstruction techniques using both numerical and experimental datasets and comparing them to an existing one-step reconstruction algorithm based on time-reversal.

II. METHODS

Iterative and non-iterative (i.e. one-step) reconstruction approaches were applied to various *in-silico*, phantom, *ex-vivo* and *in-vivo* data sets. Their performances were compared visually and evaluated quantitatively using a set of metrics.

A. Image Reconstruction Strategies

To study the performance of model-based iterative image reconstruction methods for a planar detection geometry, various non-iterative and iterative methods were used to reconstruct photoacoustic images (see Table I for abbreviations). For the non-iterative method, time-reversal (TR) was used because it uses the same propagation model as the iterative method, so any differences in the reconstructed images do not arise from differences in the propagation model but in the reconstruction approach only [30], [31], [32]. We used two variants: (i) TR with a non-negativity constraint (TR+) since physically the initial pressure distribution is not negative [33], [34] and (ii) TR with total variation de-noising [35] implemented as a post-processing step on the reconstructed image (TR+TVdenoising);

For the model-based iterative methods, the image reconstruction was posed as a regularized non-negative least-squares minimisation problem which matches the solution to the measurement (equation 1). By using this format, it provides a flexible framework for incorporating prior knowledge.

$$p_0 = \underset{p_0 \geq 0}{\operatorname{argmin}} \frac{1}{2} \|\mathcal{A}p_0 - f\|_2^2 + \lambda \mathcal{R}(p_0) \tag{1}$$

The first part of equation 1 is the data fidelity term which matches the estimate of the initial pressure distribution p_0 to the measured photoacoustic time series f. The forward operator $\mathcal{A}$ describes the mapping of the pressure distribution to the pressure time series. A more detailed discussion of the operator $\mathcal{A}$ and its adjoint can be found in [36]. The forward and adjoint models were implemented using the k-Wave Matlab Toolbox [18]. The second term in equation 1 is the regularization functional $\mathcal{R}$ which encodes a-priori knowledge about the solution. The regularization parameter λ controls the balance between the data fidelity term and the regularization functional. The final prediction of the initial pressure distribution p_0 is then achieved by iteratively minimizing the difference between the measured time series and the modelled time series.

Three model-based iterative image reconstruction methods based on equation 1 were explored.

(i) Iterative least squares (iLS+) with a non-negativity constraint but no regularization;

Thus, iLS+ is equation 1 with $\lambda \mathcal{R}(p_0) = 0$.

(ii) Iterative regularization term and a non-negativity constraint (TV+) [28];

TV+ is equation 1 with $\mathcal{R}(p_0)$ set to $TV(p_0)$ where the discrete generalization of the isotropic $TV(p_0)$ is given by as in (2), shown at the bottom of the page. $TV(p_0)$ is a measure of the L1 norm of the amplitude of the gradient field of p_0. In general, it quantifies how much an image varies across pixels. A highly textured or noisy image will have a large TV energy, while a smooth or piecewise constant image would have a relatively small TV. Hence, TV regularization encourages piecewise smooth regions yet preserves their edges while rejecting noise. TV regularization has previously been used in the reconstruction of photoacoustic images from incomplete data to suppress limited view artefacts [22], [37] and minimise image degradation [28], [38].

(iii) Bregman iteration enhanced TV (TV+Bregman);

A potential drawback of TV is that it can lead to regularization bias resulting in amplitude error and loss of fine structures [39]. To overcome this drawback, an iterative enhancement of variational solutions by the Bregman iteration [40], [41] was applied with the following reformulated form:

$$p_0^{k+1} = \underset{p_0 \geq 0}{\operatorname{argmin}} \frac{1}{2} \left\| \mathcal{A}p_0^k - (f + b^k) \right\|_2^2 + \lambda \mathcal{R}\left(p_0^k\right) \tag{3}$$

$$b^{k+1} = b^k + (f - \mathcal{A}p_0^{k+1}) \tag{4}$$

$$TV(p_0) = \sum\nolimits_{(i,j,k)} \sqrt{p_{0(i+1,j,k)} - p_{0(i,j,k)})^2 + \left(p_{0_{(i,j+1,k)}} - p_{0(i,j,k)}\right)^2 + \left(p_{0_{(i,j,k+1)}} - p_{0_{(i,j,k)}}\right)^2} \tag{2}$$

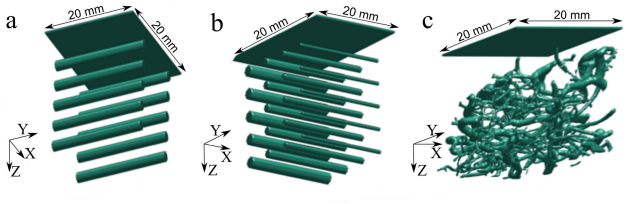

Fig. 1. Illustration of the numerical phantom geometries: a, 10-cylinder phantom; b, 20-cylinder phantom; c, vascular-like phantom. The planes on the top of each image illustrate the planar region over which the photoacoustic waves are detected.

As can be seen from equation (3) & (4), unlike TV+, the time series is iteratively updated by adding the difference between the measured and the simulated time series. By adding the difference back, the amplitude error and loss of fine structures can be mitigated. In each Bregman iteration, TV+ iteration is re-run to update the estimate of p_0.

To solve the optimisation problem (equation 1), an accelerated proximal gradient descent scheme was used because the proximal operator (here the operator that solves the 3D non-negativity constrained TV denoising problem) is cheap to compute compared to the forward $\mathcal{A}$ and adjoint $\mathcal{A}^T$ operations [28]. The basic proximal gradient scheme is represented by

$$p_0^{i+1} = \text{prox}_{TV+,\eta\lambda}\left(p_0^i - \eta\mathcal{A}^T\left(\mathcal{A}p_0^i - f\right)\right),$$
$$p_0^0 = 0, i = 1, \ldots, L$$

A FISTA-type acceleration extends this scheme [42]. The positivity-constrained TV denoising is implemented by a primal-dual hybrid gradient algorithm as described in [43]. More details about the optimisation algorithms can be found in [28].

B. Experiments

In-silico, phantom, *ex-vivo* and *in-vivo* studies were undertaken to examine the performance of the different reconstruction methods.

1) In-Silico Studies: To evaluate the different image reconstruction methods, a range of 3D numerical phantoms comprising cylinders of different dimensions and vascular-like structures were used in conjunction with an acoustic propagation model to generate photoacoustic time series data.

As shown in Fig1a-c, three phantom geometries were used: 10-cylinder phantom, 20-cylinder phantom and vascular-like phantom. In the 10-cylinder phantom, ten cylindrical targets with identical radii of 0.4 mm were equally spaced in depth. The 20-cylinder phantom comprised twenty equally spaced cylinders with different radii of 0.4, 0.3, 0.2 and 0.1 mm. The vascular-like phantom was based on a segmentation of a micro-CT scan of a rabbit lymph node. The initial pressure p_0 was set to be constant within the cylinders (or vessels) and zero elsewhere.

A variation of the 10-cylinder phantom in which a non-uniform p_0 distribution was produced was also used and is referred to as the non-constant 10-cylinder phantom in this

TABLE II
OPTICAL PROPERTIES IN LIGHT TRANSPORT MODELLING

Optical properties	Background	Tube
Absorption coefficient (mm^{-1})	0.001	0.4049
Scattering coefficient (mm^{-1})	12.3965	272.49
Anisotropy	0.9	0.994
Refractive index	1.376	1.36

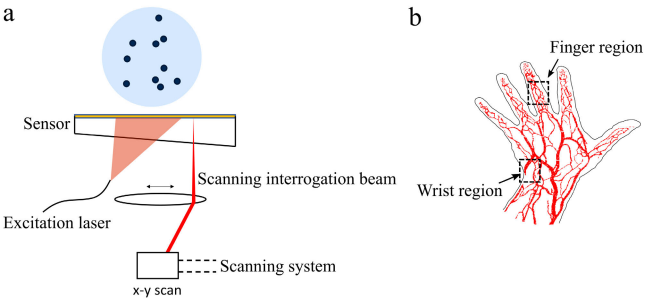

Fig. 2. Experimental set-up. a, FP scanner and wall-less agar phantom; b, In-vivo imaging of human finger joint and wrist.

paper. This was formed by using a Monte Carlo model of light transport (MCXLAB [44]) to simulate the fluence distribution in the phantom. The optical coefficients of the cylinders and the background are given in (Table II) [45]. The optical properties of the cylinders were set to those of blood. The background was mimicking a capillary bed, modelled as 55% water and 45% blood with a total haemoglobin concentration, c_{HbT}^{bg} =5.6gl^{-1}3 and 60.7% oxygenation (c_{HbO_2} =3.42gl^{-1} and c_{Hb} =2.11gl^{-1}). All optical coefficients were chosen to be those at a wavelength of 784 nm. The volume was illuminated from the top with a collimated Gaussian beam with 5 mm waist radius. The fluence distribution was then multiplied pixel-wise by an image of the corresponding optical absorption coefficient to produce images of the absorbed energy distribution and thus p_0 (assuming a Grueneisen coefficient of one).

The same simulation grid consisting of 100×100×100 cubic voxels of length 20 mm was used in all the simulation studies and the speed of sound was assumed to be homogeneous at 1500 m/s. 2D planar sensor arrays with identical size of 20 × 20 mm were positioned at the top of each simulation grids to acquire the time series data. The distance between detector elements was 200 μm, the same as the grid size. The temporal sampling interval was set to be 40 ns. The sensor bandwidth was not explicitly limited but due to the grid size, the bandwidth of the photoacoustic signals is limited to 5.5 MHz. The photoacoustic wave propagation from the initial pressure distribution and its subsequent detection by the 2D planar array was simulated using k-Wave [18]. Gaussian noise was added to the detected pressure time series to achieve a range of SNRs from 10dB to -10dB.

2) Experimental Studies: Measured photoacoustic data was obtained in a phantom (Fig. 2), ex vivo tissue and in vivo using a Fabry-Pérot (FP) planar scanner [46]. Fig 2a shows a schematic of the scanner. The excitation light is emitted by a fibre coupled, Nd:YAG laser in the phantom experiment or a OPO laser system in ex-vivo and in-vivo experiments. The light is transmitted through the sensor head, which is designed

to be transparent to the excitation wavelengths. The absorption of the excitation light generates photoacoustic waves, which propagate to the FP sensor and modulate its optical thickness and thus its reflectivity. The latter is detected using a focused interrogation laser beam, that is raster scanned over the sensor surface. The interrogation beam is reflected from the sensor and is incident on a photodiode, which is connected to a digitizer to record the photoacoustic waveforms. The FP sensor provides a near-uniform broadband frequency response from 50 kHz to 22 MHz (-3dB). The lateral resolution of the FP scanner is spatially variant. It increases with depth and the horizontal distance from the centre of the scan. The lateral resolution is in the range of 50 to 125 μm and the axial resolution is 27 μm. The scanner was used to image a phantom comprising a pure agar background containing 10 absorbing wall-less cylindrical agar inclusions $\mu_a = 2mm^{-1}$ of diameter 1 mm. The agar background has negligible optical absorption and optical scattering [47]. The phantom was illuminated with a 3 cm beam diameter at an excitation wavelength of 1064nm. An area of 14 mm $\times$ 16 mm was scanned, with a step size of 120 μm, to yield $\approx$16,000 time series. Each time series contained 1200 time points, with a temporal resolution of 20 ns. Photoacoustic signals were initially acquired with the phantom in a single fixed position in order to reconstruct a limited view image. Thereafter, the phantom was rotated enabling signals to be acquired from 6 planes equally distributed around the phantom. The sensor coordinates were obtained from a registration procedure [48]. The combined time series from all 6 planes were then used to reconstruct a full view image for comparison with the single view image.

An image of a human placenta was acquired at a single plane using the FP planar scanner. The tissue was illuminated with a 6 mm diameter laser beam at an excitation wavelength of 850nm. An area of 8.2 mm $\times$ 8.1 mm was scanned, with a step size of 100 μm. Each time series contained 1000 timepoints with a temporal resolution of 20 ns.

The skin vasculature in the finger and wrist region of a volunteer were imaged at a single plane using the FP planar scanner (as shown in Fig 2b) [49]. The data was acquired with an excitation wavelength of 850 nm. An area of 15.26 mm $\times$ 14.95 mm was scanned, with a step size of 106 μm. Each time series contained 500 time points with a temporal resolution of 16.67 ns. Ethical permission was granted for this prospective single-centre study by the University College London Research Ethics Committee (Project ID: 1133/001).

C. Quantitative Metrics

In order to quantitatively compare the images reconstructed by the different methods, the following metrics were variously used. In the simulation studies, mean squared error (MSE) was calculated to evaluate the extent to which the reconstructed image differs from the ground truth. MSE was estimated from the reconstructed solution p and the ground truth phantom p_0 as follows:

$$MSE = \frac{1}{N}\sum_{i}^{N}(p_i - p_{0_i})_2^2 \quad (5)$$

Signal-to-noise ratio (SNR) was estimated to account for the random noise generated by the acoustic detection instrumentation that is mapped on to the image. SNR is defined as the ratio of the mean intensities of a reconstructed image feature and a region of random noise selected such that it does not overlap with any reconstructed features and their associated artefacts:

$$SNR = 20\log\left(\frac{mean\ intensity\ over\ feature\ region}{mean\ intensity\ over\ noise\ region}\right). \quad (6)$$

Additionally, signal-to-artefact ratio (SAR) was evaluated to provide a measure of the level of artefacts in the image. SAR was defined as the ratio between the mean intensities over the feature region and the artefact region:

$$SAR = 20\log\left(\frac{mean\ intensity\ over\ feature\ region}{mean\ intensity\ over\ artefact\ region}\right) \quad (7)$$

The methods used to identify regions of feature, artefact and noise differed depending on whether the images were reconstructed using simulated or experimental data. Fig 3 illustrates the different approaches. Fig 3 a&b show the cross-sectional image of the initial pressure distribution p_0 and the corresponding TR+ reconstructed image. For the in-silico studies, K-means clustering image segmentation was applied to identify the feature, artefact and noise regions. As shown in Fig 3c, the dashed contour indicates the feature region segmented by K-means method. In Fig 3d&e, TR reconstructed images are overlaid with masks of artefact and noise ROIs detected by K-means segmentation [50]. For the images of the experimental phantom, feature regions were segmented similarly by the K-means method while artefact and noise regions were manually selected. For ex-vivo and in-vivo experiments, the feature, artefact and noise regions were manually identified. An example of the manual selection of artefact and noise regions is illustrated in Fig 3f.

The Hausdorff distance was used to evaluate the geometrical distortion, the difference between a distorted feature and the ground truth [51]. A smaller Hausdorff distance means the reconstructed geometry is closer to the ground truth.

III. Results

A. In-Silico Results

1) 10-Cylinder Phantom: The numerical phantom composed of 10 identical cylinders of uniform p_0 was first used to compare the non-iterative and iterative reconstruction methods. Fig 4a-e shows the cross-sectional images reconstructed from simulated pressure time series with an SNR of 5dB. As Fig 4b shows, the image reconstructed using TR+ exhibits significant streaking artefacts around the cylinders. The latter are also distorted, and the amplitudes of the features are far from the true values. These distortion and amplitude errors increase with increasing depth. TVdenoising was applied to non-iterative TR+ to determine whether TV+ achieves more than just smoothing out the noise which TVdenoising can achieve with similar effectiveness. As Fig 4c shows, the noise and the streaking artefacts apparent in TR+ are partially reduced

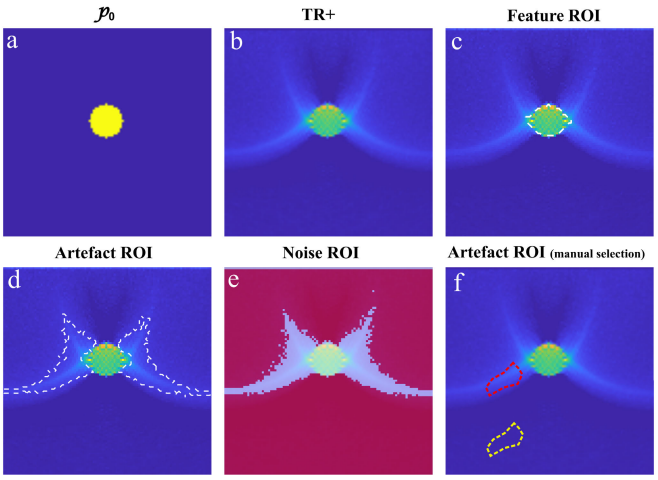

Fig. 3. Illustration of image segmentation to quantify artefact and noise level: a, 1-cylinder phantom; b, TR+ reconstructed image; c, feature ROI (K-means segmented) indicated by white dashed contour; d, TR+ image overlaid with boundaries of artefact ROI (K-means segmented); e, TR+ image overlaid with red mask of noise ROI (K-means segmented); f, manually selected ROIs for artefact (red) and noise (yellow).

by TVdenoising. However, the geometrical distortion and the reduction in amplitude with depth remain. Both are mitigated by the two iterative methods., iLS+ exhibits higher noise and greater artefacts than TV+ ($\lambda = 10^{-1}$) which supresses the streaking artefacts and reduces the geometrical distortion.

Fig 4f-k provides a quantitative comparison of images of the 10-cylinder phantom reconstructed using pressure time-series data of different SNRs. As shown by the MSE (Fig 4f), in contrast to TR+ and TR+TVdenoisng, the reconstructions achieved by the iterative methods (iLS+ and TV+) are closer to the ground truth for time series SNRs in the range 10dB to −5dB. It is only at the lowest time series SNR (−10dB) that TR+TVdenoising provided lower MSE than iLS+. TV+ outperforms the other methods and achieves the lowest MSE at all time series SNRs. To permit a comparison of the amplitude accuracy, profiles along the vertical and horizontal dashed lines shown in Fig 4a are plotted in Fig 4g&h. This plot shows that, compared to the non-iterative methods, the two iterative methods (iLS+ and TV+) achieve higher amplitude accuracy, although both still tend to under-estimate the amplitude. For TV+, this is likely to be due to the TV regularization bias and is discussed further in the following section (Enhancement via Bregman iteration). As shown Fig 4i, iterative methods provide higher SAR than the non-iterative methods for the SNRs in the range 10dB to −5dB reflecting higher artefact suppression. Only at the lowest time series SNR (−10dB) did TR+TVdenoising achieve higher SAR than iLS+. However, TV+ outperforms the other methods and obtains the highest SAR at all time series SNRs.

The Hausdorff distances of the features indicated by the white dashed boxes in Fig 4a were calculated to evaluate the geometrical distortion introduced by the different reconstruction methods. As we can see from Fig 4d & 4e, the correction of geometrical distortion is less effective with increasing depth.

This is illustrated quantitatively by the Hausdorff distances in Fig 4j&k. Fig 4j corresponds to the feature in the second row and shows that the Hausdorff distance of the iterative reconstruction methods (iLS+ and TV+) are zero at all time series SNRs. Fig 4k corresponds to the deeper lying feature in the lowest row and suggests there is no clear advantage of iLS+ over non-iterative methods. TV+ however, outperforms the other methods, although is increasingly less effective as the SNR decreases.

2) Impact of Regularization Parameter: For TV+, the regularization parameter λ controls the balance between the data fidelity term and the regularization functional and thus has a significant impact on image fidelity. Small λ values can result in a reconstruction result similar to that achieved by iLS+ where noise and artefacts remain. On the other hand, if λ is too large, fine structures can be lost and amplitude accuracy compromised. To explore this, three numerical phantoms were used to evaluate the impact of the regularization parameter on image fidelity and contrast as a function of the target geometry and intensity distribution. In Fig 5, images reconstructed using TV+ for four different λ values ($\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$) are compared qualitatively and quantitatively.

The first numerical phantom was the 10-cylinder phantom where each cylinder is assigned the same p_0. The corresponding reconstructed cross-sectional images are shown in the first subgroup (i) of Fig 5(Fig 5(i)a-e). This shows that the suppression of noise and artefacts as well as the correction of distortion are less effective for smaller λ values. Fig 5(i)f-h show the corresponding MSE, SAR and Hausdorff distance. These results show that the lowest MSE is achieved with the largest λ value (Fig 5(i)f) while the SAR reduces with decreasing λ (Fig 5(i)g) due to the reduced artefact suppression associated with small λ. The feature used to calculate the Hausdorff distance is indicated by the white dashed box in Fig 5 (i)a. The higher level of geometrical distortion produced by λ_3 & λ_4 is reflected by the larger Hausdorff distance values in Fig 5(i)h.

In the above simple 10-cylinder phantom where the image features are identical and of large dimensions, the use of a large λ is optimal and a relatively straightforward choice. However, in images comprising different sized features the selection of λ can be more nuanced and challenging. This is because TV can suppress small features on account of their relatively higher edge-densities which result in higher TV energy [52]. When the TV energy of the entire image is minimised, such features are at higher risk of partial or complete removal by regularization while large-scale features remain relatively unaffected.

To assess the scale-dependent nature of TV regularization [52], the effect of varying λ was studied using the 20-cylinder phantom which comprises differently sized cylinders. The reconstructed images are shown in the second subgroup (ii) of Figs 5 (Fig 5(ii)a-e). As λ increases, targets of smaller size or at greater depth are poorly reconstructed or missing altogether. As Fig 5(ii)b (largest λ value) shows, if we compare cylinders at the same depth, targets of smaller size are poorly reconstructed or missing, a consequence of the above mentioned scale-dependent property of TV regularization. For cylinders of the same size and at different depths (each column), those at greater depth were poorly reconstructed or missing due

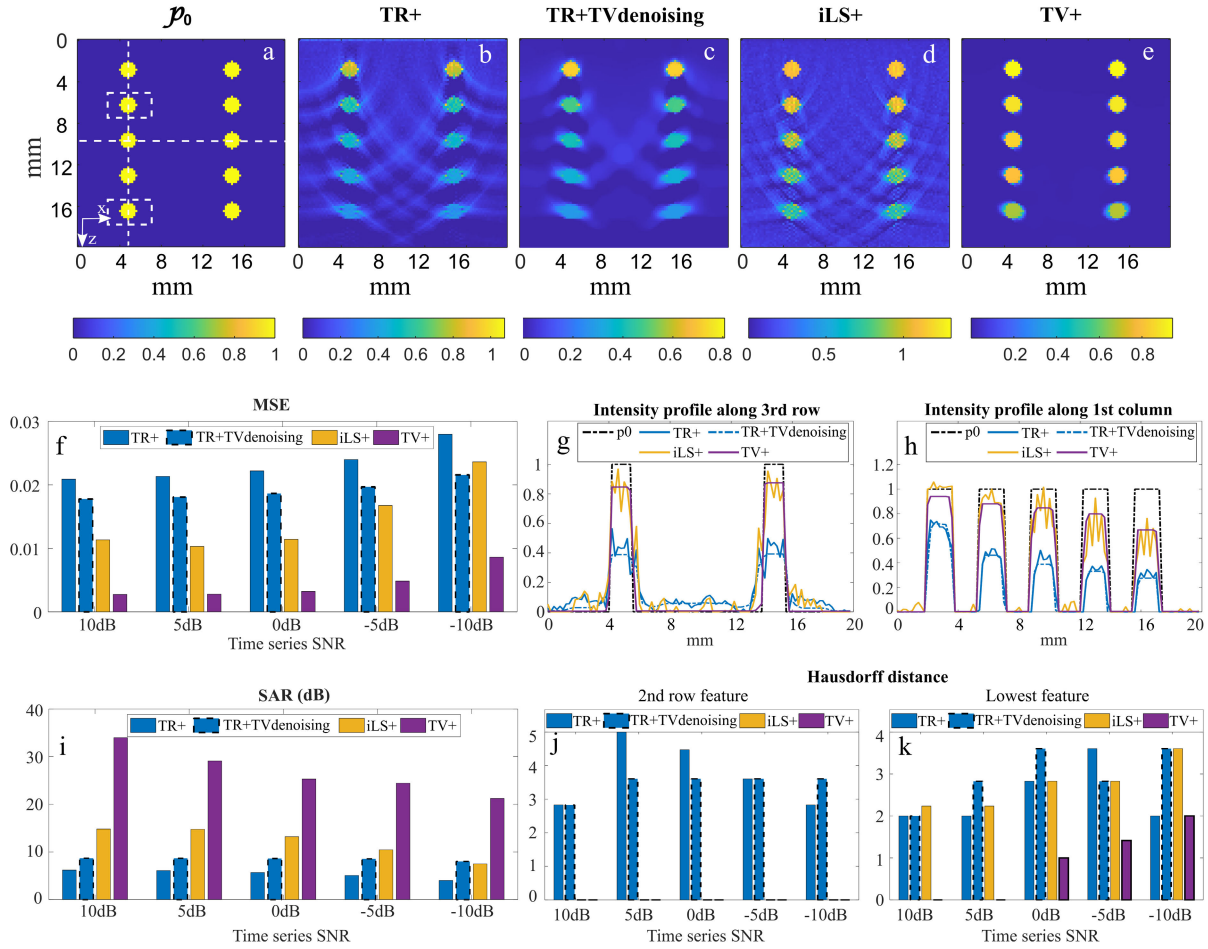

Fig. 4. Cross-sectional single-slice image of the 10-cylinder phantom: a, phantom p0; b, TR+; c, TR+TVdenoising; d, iLS+; e, TV+; Quantitative comparison of reconstructed images achieved by non-iterative and iterative methods for different pressure time series SNRs: f, mean squared error (MSE); g, h, intensity profiles; i, SAR; j, k, Hausdorff distance.

to the lower SNR associated with them. The MSE values (Fig 5(ii)f) suggest that $\lambda_3$is the optimal regularization parameter value for this phantom. This is further evidenced by the amplitude profiles along the dashed lines (Fig 5(ii)a) shown in Fig 5(ii)g & h which are the most accurate for λ_3. By contrast the larger λ_1 results in significant amplitude errors and the smaller cylinders and cylinders at greater depth are invisible. λ_4 enables all features to be visualised but yielded the highest noise as well as lower amplitude accuracy than λ_3. These results show that choosing a value of λ that is too large results in the loss of small features. However, selecting λ that is too small results not only in more noise but greater amplitude errors. Hence there is a compromise to be made to find the best trade-off between small feature visibility and amplitude accuracy.

Furthermore, it is known that TV encourages piece-wise smooth regions. The 10-cylinder and 20-cylinder phantoms, where the image features have a uniform amplitude distribution, may not therefore be a sufficient test since real photoacoustic images are rarely piece-wise constant, not least because of the spatially varying light fluence. Under these conditions, there is risk of TV regularization biasing the reconstruction towards constant p_0 and reducing image amplitude accuracy. To investigate this, the non-constant 10-cylinder phantom described in section B was used; with this phantom the light fluence is simulated resulting in a non-uniform p_0 both within each cylinder and the background. Fig 5(iii)a shows the p_0 distribution. Since the volume was illuminated from the top, the light fluence decreases with depth. Hence, a depth dependent decrease in p_0 within each cylinder can be observed. Fig 5(iii)a-e shows the corresponding cross-sectional images obtained for different λ values. For the smallest λ used (λ_4), the image shown in Fig 5(iii)e and the profiles in Figs 5(iii)g & h suggest the amplitude accuracy is high, although for the deeper cylinders there is significant noise and evidence of streaking artefacts. As λ increases to $\lambda_3$the amplitude accuracy increases and the noise decreases, as evidenced not only by the profiles in Fig 5(iii)g & h but also by the MSE which is the lowest for all λ (fig 5(iii)f). For higher λ, the noise and streaking artefacts decrease further but at the cost of amplitude accuracy and the visibility of deeper lying features. For example, for the highest λ value (λ_1), the top row of cylinders are almost piecewise constant and thus do not reflect the variation of p_0 within each cylinder while the deeper

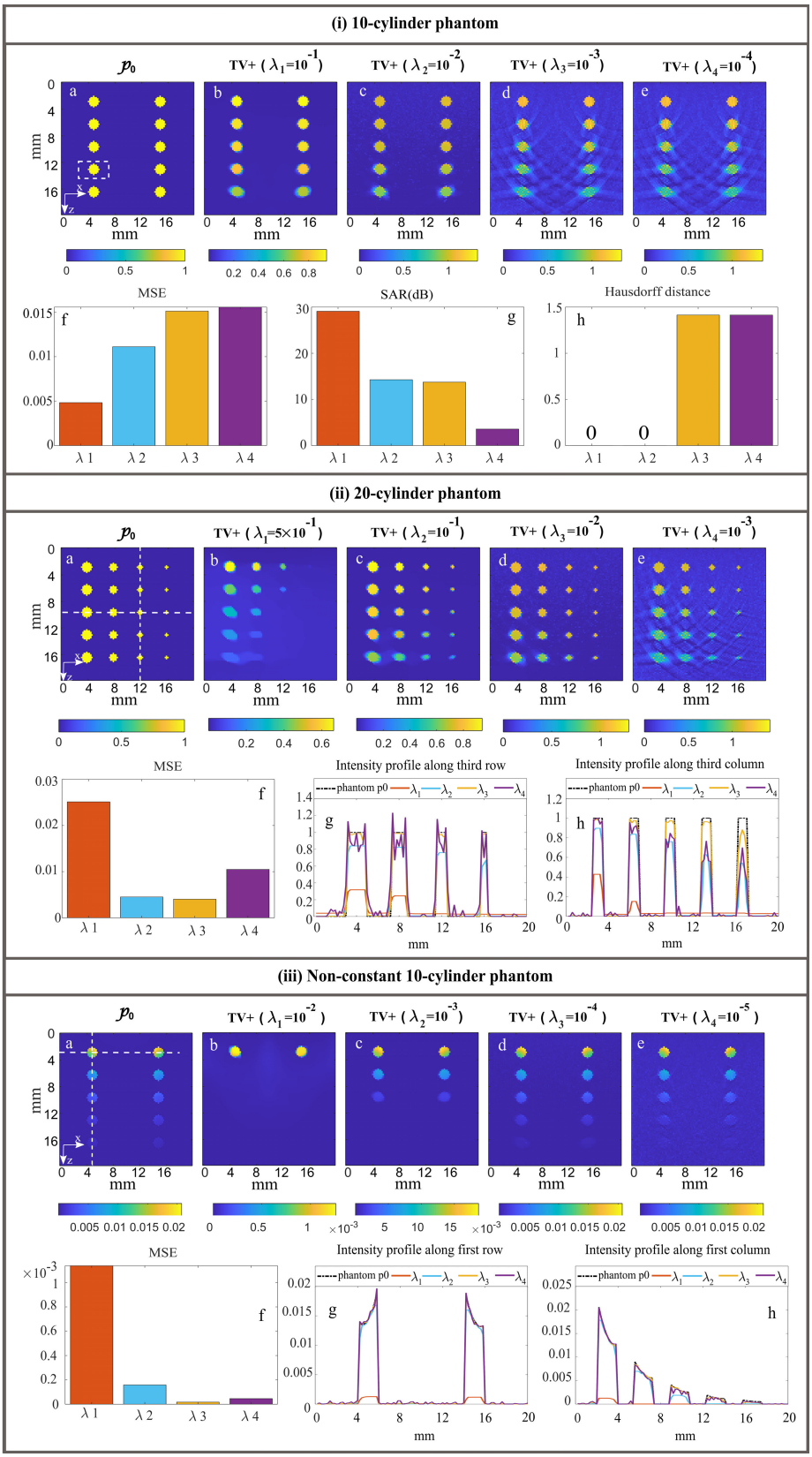

Fig. 5. Comparison of images reconstructed using TV+ for different regularization parameters: (i), 10-cylinder phantom $\lambda_1 = 10^{-1}$, $\lambda_2 = 10^{-2}$, $\lambda_3 = 10^{-3}$, $\lambda_4 = 10^{-4}$; (ii), 20-cylinder phantom $\lambda_1 = 5 \times 10^{-1}$, $\lambda_2 = 10^{-1}$, $\lambda_3 = 10^{-2}$, $\lambda_4 = 10^{-3}$; (iii), Non-constant 10-cylinder $\lambda_1 = 10^{-2}$, $\lambda_2 = 10^{-3}$, $\lambda_3 = 10^{-4}$, $\lambda_4 = 10^{-5}$.

lying features are invisible. For this example, λ_3 appears to provide the best compromise in terms of noise reduction, artefact suppression and amplitude accuracy. Notably, these results show that, although TV promotes a piece-wise constant

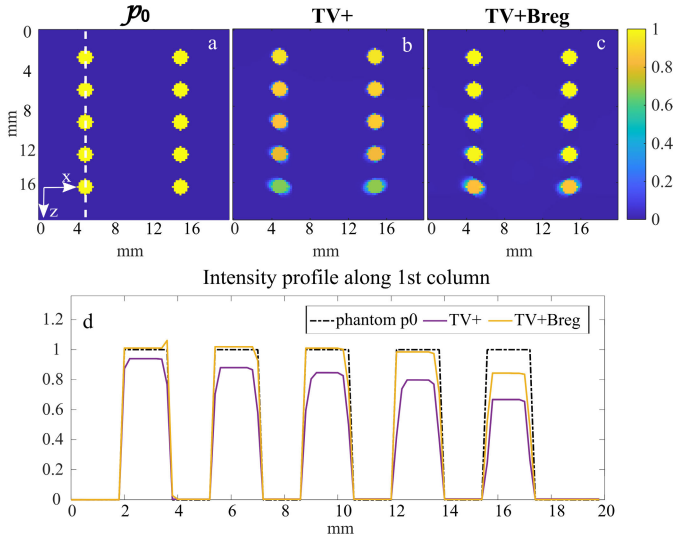

Fig. 6. Total variation regularization enhanced by Bregman iteration: a-c, visual comparison of cross-sectional reconstructed images; d, intensity profile along 1st column. $\lambda_{\text{TV+}} = 10^{-1}$, $\lambda_{\text{TV+Breg}} = 5 \times 10^{-1}$.

solution, a realistic non-uniform p_0 distribution can still be accurately recovered with a carefully selected regularization parameter.

3) Enhancement via Bregman Iteration: The amplitude errors (observed in Fig 4g & h) introduced by TV regularization can be compensated through the use of Bregman iteration (equations (3) & (4)). The effectiveness of this approach was investigated using the 10-cylinder phantom with uniform pressure distribution. Fig 6 compares the results achieved by TV+ and TV+Breg. As illustrated by Figs 6a-c, Bregman iterations act to reduce the amplitude errors due to TV regularization. This further evidenced by the profiles along the dashed line in Fig 6a which are plotted in Fig 6d. The images reconstructed with Bregman iterations achieved the most accurate amplitude, although the improvement decreases at the greater depth. A total of 5 Bregman iterations were performed in this and all the following experiments.

4) Vascular-Like Phantom: In the previous simulation studies, numerical phantoms composed of simple cylindrical objects were used. Although they can provide insights in the basic properties of reconstruction methods, it is often unclear how their results will translate to more complex targets. Hence, the vascular-like phantom described in section IIB(1) was used to study how the different image reconstruction methods are likely to perform when imaging blood vessels. Fig 7a-f show the depth colour-coded x-y *en-face*MIP images. In the images reconstructed by non-iterative methods (TR+ and TR+TVdenoising), only large vessels are clearly visualised while the iterative methods (iLS+, TV+ and TV+Bregman) permit visualisation of smaller vessels. Fig 7g-l shows cross-sectional x-z single-slice images. In the images reconstructed by TR+, TR+TVdenoising and iLS+, the true vascular structure is not visualised due to the artefacts. TV+ and TV+Bregman provides images with significantly fewer artefacts and thus a vascular structure that corresponds more closely to the ground truth. Furthermore, the size of the vessel

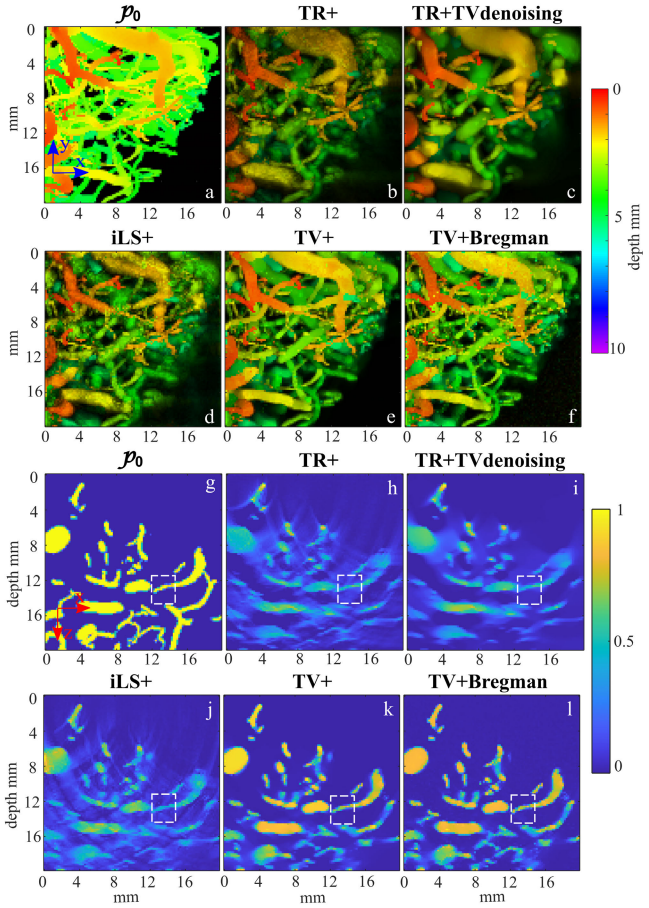

Fig. 7. Comparison of non-iterative and iterative reconstruction methods using vascular-like phantom: a-f, x-y MIP; g-l, cross-sectional x-z single slice images. $\lambda_{\text{TV+}} = 10^{-1}$, $\lambda_{\text{TV+Breg}} = 10^{-1}$. White dotted rectangle indicates vessel selected for size measurement provided in Table III.

TABLE III
FULL WIDTH HALF MAXIMUM (FWHM) VALUE MEASURED THE SIZE OF THE VESSEL IN THE WHITE DASHED BOX IN FIG 7 g-l

Phantom	**TR+**	**TR+TVdenoising**
0.273 mm	0.385 mm	0.327 mm
iLS+	**TV+**	**TV+Breg**
0.338 mm	0.278 mm	0.274 mm

in the white dashed box was quantified by the full width half maximum (FWHM). As we can see from Table III, TV+Bregman achieves the most accurate FWHM.

B. Experimental Phantom

To evaluate the reconstruction methods using experimental data, the wall-less agar phantom containing 10 absorbing cylindrical inclusions each of diameter 1 mm shown in Fig2a was used. The phantom is located in front of the sensor and rotated by a stepper motor to obtain photoacoustic time series data over 6 different planes as shown in Fig 8. A full-view image was then reconstructed using time-reversal from this data. This image does not suffer from limited view artefacts. Hence, it is taken to represent the ground truth but only in terms of the image structure – the amplitude

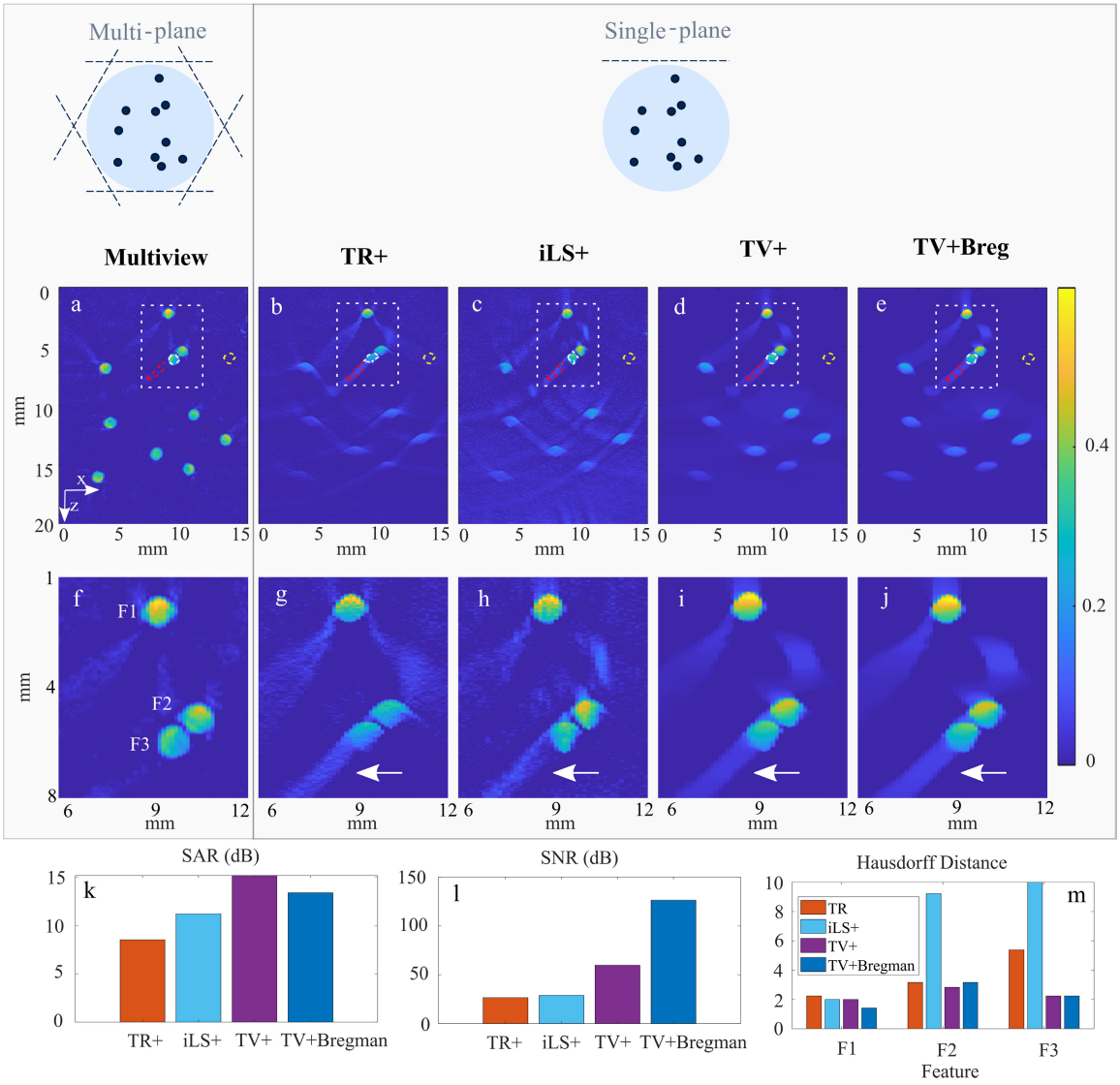

Fig. 8. Cross-sectional single-slice images of the agar phantom. a-e, single slice images with ROI contours: white dashed boxes indicated the three features used to compare hausdorff distance. In the white dashed boxes, red dashed contours indicate artefact region, white dashed contours indicate feature region. Outside the white dashed boxes, yellow contours indicate noise region; f-j, corresponding zoomed-in images of the white dashed boxes in a-e; k, SAR comparison; l, SNR comparison; m, Hausdorff distance comparison. $\lambda_{TV+} = 10^{-2}$, $\lambda_{TV+Breg} = 5 \times 10^{-2}$.

distribution is subject to errors because the illumination of the phantom is not the same for each plane. Thereafter images were reconstructed from photoacoustic time series data acquired over a single plane using non-iterative and iterative methods. The regularization parameter λ was chosen to be large enough to suppress most visible artefacts and the most visually apparent geometrical distortion without compromising the visualisation of low SNR features. Fig 8a-e show the cross-sectional single-slice images of the entire phantom while Fig 8f-j show the zoomed-in images of the top three features within the white dashed boxes. Compared to TR+, iLS+ enhances the visibility of deeper targets but the artefacts and geometrical distortion remain. TV+ and TV+Bregman suppress the artefacts and reduce the geometrical distortion, albeit with reduced effectiveness with increasing depth. From the zoomed-in images (Fig 8g-j) it is apparent that the large streaking artefact region indicated by the white arrow has been transformed to a relatively high amplitude uniform feature by TV regularization. This is because the artefact region is of similar scale as the image feature from which it originates. Furthermore, the artefact region has higher intensity than the image features at deeper depth. If a larger regularization parameter is used to suppress the artefact, image features at deeper depth will be removed as well.

The SAR and SNR were measured within the ROIs indicated in Fig 8a-e. The white, red and yellow dotted line contours represent the feature, artefact and noise ROIs respectively. As shown in Fig 8k, TV+ achieves the best SAR as it most effectively suppresses the artefacts whereas TV+Bregman provides highest SNR (Fig 8l). The Hausdorff distance for the top three features was calculated using the full-view image as the reference image. As can be seen from Fig 8m, TV+ and TV+Bregman provide more accurate target geometry while iLS+ achieves the least accurate target

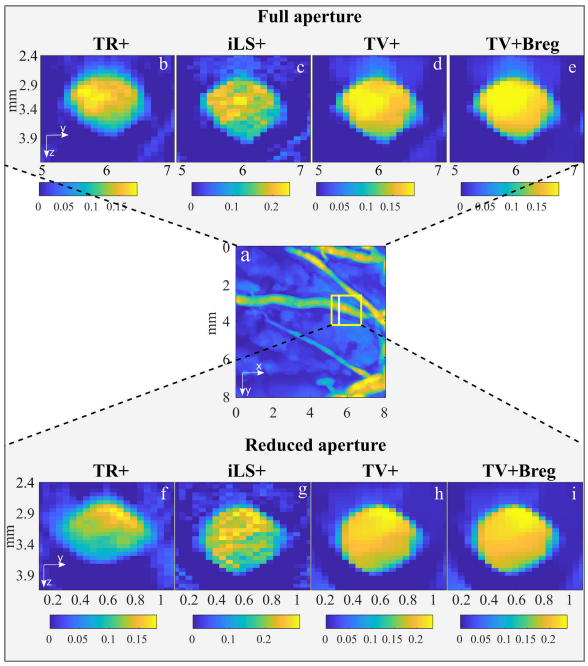

Fig. 9. Ex-vivo placenta images; a, x-y MIP image reconstructed from signals acquired over the full aperture (8 × 8 mm), b-e, y-z cross-sectional single-slice full aperture images of region indicated by white line in (a), f-i, corresponding y-z images reconstructed from signals acquired over the reduced 2 × 2 mm aperture indicated by yellow rectangle in (a). $\lambda_{TV+} = 10^{-3}$, $\lambda_{TV+Breg} = 5.5 \times 10^{-3}$.

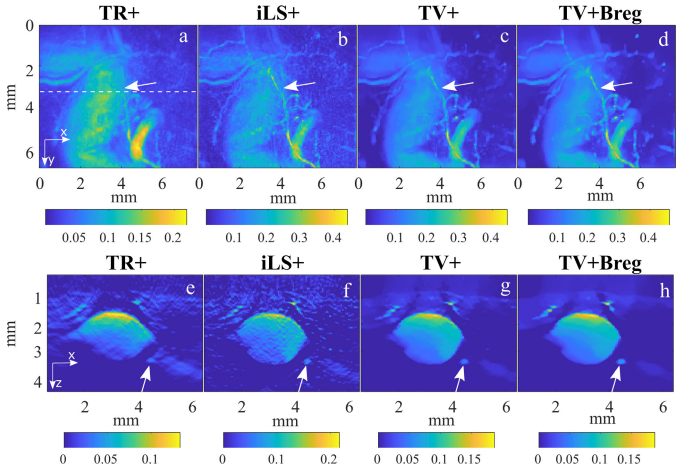

Fig. 10. In-vivo human finger joint images: a-d, x-y MIP images reconstructed by non-iterative and iterative methods; e-h, cross-sectional single-slice images reconstructed by non-iterative and iterative methods. The arrows indicate the same small vessel. The dotted white line in (a) indicates the location of the single-slice images e-h. $\lambda_{TV+} = 10^{-3}$, $\lambda_{TV+Breg} = 5.5 \times 10^{-3}$.

geometry. MSE and intensity profiles are not provided. As mentioned above this is because, for the full-view image, the ground truth amplitude distribution in the cylindrical inclusions is unknown.

C. Ex-Vivo and In-Vivo Results

The performance of the iterative reconstruction methods was further evaluated by acquiring measured data from an *ex vivo* human placenta and the superficial vasculature in the finger and wrist of a human volunteer. There is no absolute ground truth in these examples. However, they can still provide an insight into the effectiveness of the different reconstruction methods if reasonable assumptions are made about the tissue anatomy; for example, that vessels are near cylindrical in shape. To quantify the suppression of noise and artefacts, metrics such as SNR and SAR can still be used to evaluate performance.

From the simulation and experimental phantom results, it is clear that the regularization parameter λ must be appropriately chosen. This was performed by selecting a value of λ that was large enough to 1) suppress most visible reconstruction artefacts 2) retain most visible small vessels and 3) correct the most obvious geometrical distortion of clearly visualised vessels.

Fig 9 shows the ex-vivo human placenta images. Fig 9a shows the image reconstructed by TR+ using the photoacoustic signals acquired over the full detection aperture of 8 mm x 8 mm. Signals from a reduced aperture of 2mm x 2mm (indicated by the yellow box in Fig 9a) were then extracted from the full aperture data set. This was done because, when using the full detection aperture, the image distortion and artefact level due to limited view was found to be modest, a consequence of the superficial nature of the blood vessels in the tissue sample used. By artificially reducing the detection aperture, the limited view problem is exacerbated. The artefacts therefore become more severe and thus the extent to which they are reduced by the iterative methods more obvious.

Fig 9b-e show cross-sectional single-slice images of a blood vessel taken from the full aperture while Fig 9f-i show images of the same region from the reduced aperture image. The white line in Fig 9a indicates the location of the single slice. As we can see from the comparison between Fig 9b and Fig 9f, the TR+ image obtained from the reduced aperture exhibits more distortion than the full aperture image. In both cases the iterative reconstruction methods reduce the distortion although it is more apparent in the reduced aperture images.

The in-vivo images of the finger vasculature are shown in Fig 10. Fig 10a-d show the x-y MIP images. The small vessel indicated by the white arrow can be clearly visualised in the images reconstructed by iterative methods (iLS+, TV+ and TV+Bregman), while the same vessel is poorly visualised in TR+. Fig 10e-h show cross-sectional single-slice images extracted from the reconstructed 3D image at the location indicated by the dotted horizontal white line in Fig 10a. For the large vessel in the middle of the single slice images, the iterative methods (iLS+, TV+ and TV+Bregman) correct the geometrical distortion apparent in the TR+ image. Additionally, the visibility of the small vessel indicated by the white arrows is enhanced by the TV+ and TV+Bregman methods.

Fig 9 and Fig 10 demonstrate the advantages of iterative reconstruction methods in the correction of geometrical distortion. Fig 11 shows the performance of the iterative reconstruction methods in artefact reduction and enhancement. In the TR+ image (Fig 11a), the streaking artefacts are indicated by the red arrow. As can be seen from Fig 11c&d,

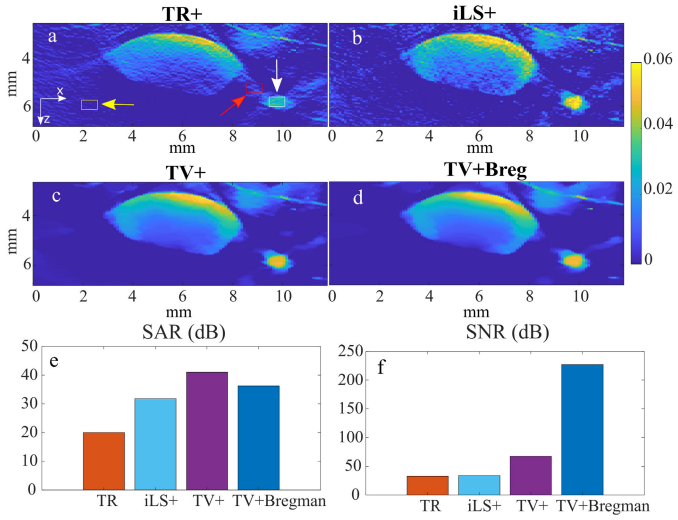

Fig. 11. In-vivo human wrist images; a-d, cross-sectional single-slice images achieved by non-iterative and iterative methods; ROI indication, white, red and yellow arrows and boxes indicate the feature, artefact and noise ROIs; e, SAR comparison; f, signal-noise-ratio comparison. $\lambda_{TV+} = 10^{-3}$, $\lambda_{TV+Breg} = 5.5 \times 10^{-3}$.

TV+ and TV+Bregman significantly remove the artefacts. SAR and SNR were measured using the ROIs indicated in Fig 11a. The white, red and yellow arrows and line contours represent the feature, artefact and noise ROIs respectively. As can be seen from Fig 11e& f, the results agree with that of the phantom experimental results. TV+ achieves the best SAR while TV+Bregman provides the highest SNR.

IV. Discussion and Conclusion

In conclusion, we have investigated the application of model-based iterative reconstruction methods in photoacoustic tomography for a planar detection geometry. The performance was assessed objectively and quantitatively using simulated and experimental datasets. Furthermore, the impact of regularization was studied and the compensation of TV regularization bias by Bregman iteration was evaluated. It has been demonstrated that the iterative reconstruction methods can provide better image quality and suppress artefacts due to the incomplete data acquired using a finite detection aperture. Amongst the iterative methods, iLS+ is unregularized so avoids the challenges involved in selecting the regularization parameter λ. However, it is least effective in reducing noise and artefacts as well as the correction of geometrical distortion. TV+ achieves better SAR by suppressing limited view artefacts while TV+Bregman achieves more accurate amplitude accuracy and better SNR but incurs additional computational cost. Both require regularization and thus involve the non-trivial task of selecting an appropriate λ as discussed below. The choice of iterative reconstruction method for any particular situation will therefore depend on requirements. iLS+ lends itself to scenarios where few assumptions about the imaging target can be made. TV+ is applicable where spatial fidelity is key and sufficient prior information exists to make an informed choice of λ. The higher amplitude accuracy of TV+Bregman could lend it to quantitative photoacoustic imaging [53] where an accurate representation of the fluence-encoded image contrast distribution is an important requirement.

Appropriate choice of λ is clearly key to the effective use of TV+ and TV+Bregman. In the absence of a ground truth, as is invariably the case in practice, this presents a challenge. As evidenced by Fig 5, when λ is too high, the scale-dependent and piecewise constant nature of TV regularization means smaller features or subtle variations in p_0 can be removed. On the other hand, when λ is set too small, the artefacts and noise remain high, and the geometrical distortion cannot be effectively corrected. The challenge becomes particularly acute when there is significant geometrical and SNR diversity in the image as the trade-off between retaining small feature visibility, minimising noise and maximising amplitude accuracy becomes finely balanced. Knowledge of the expected imaging scenario can be used to pragmatically inform the selection of λ. For example, if the anticipated target is a vascularised tumour, *in-silico* studies using numerical phantoms with tumour-realistic vascular architectures can be used to identify a regularization parameter space from which a hand-crafted value of λ can be selected for a specific *in vivo* data set. The demands of the specific application will further inform the choice of λ. If suppressing artefacts and minimizing spatial distortion is important in order to achieve high structural accuracy, a relatively large λ might be used. However, if high amplitude accuracy is the priority, as is the case for quantitative imaging, selecting a smaller λ may be prudent to avoid the promotion of an inappropriately piece-wise constant image, albeit at the cost of increased noise. Although this somewhat *ad hoc* approach has the advantages of simplicity and pragmatism, it suffers from several limitations. Firstly, it relies on the use of a global regularization parameter which will compromise image quality if there is a range of differently sized image features. To address this, a scheme could be developed to spatially adapt the regularization parameter [54]. Secondly, it requires an element of prior knowledge which may not always be available and the hand-crafting element associated with selecting λ is subjective to some extent. More objective selection of λ without the need for priors could potentially be achieved using selection methods based on the statistical properties of the data [55].

The model-based framework discussed in our study can incorporate frequency and directional responses of arbitrary complexity. While we focused on broadband, omni-directional detectors in this work, an interesting extension would be to consider more realistic detector models, e.g., directional, narrow-band ones and see how this affect the choice of regularization parameter. In the specific case of narrowband detectors with a peaked frequency response, TV regularization could mitigate reconstruction artefacts, such as the loss of the contrast in the interior of objects as well as limited view.

It has been shown that the iterative model-based approach described in this study can accommodate the large data sets ($>10^6$) associated with high resolution 3D PAT. In principle, iterative matrix-based model inversion schemes could also be used but at present are intractable for the large number of voxels encountered in high resolution 3D imaging. They are better suited to the much smaller data sets ($<10^3$) associated

with 2D PAT and have been used with the aim of mitigating the limited view problem for a linear array geometry [27]. For this geometry they may be faster than the current approach, albeit at the cost of flexibility since the model matrix has to be pre-computed for a specific set of imaging parameters. If the latter change, then matrix has to be re-computed at significant computational cost.

A limitation of iterative model-based reconstruction methods is that they are typically computationally intensive, especially for 3D image reconstruction. For example, computing the reconstruction of the finger joint images in Fig 10(486∗71∗71 voxels) took 1h18m using TV+ (50 iterations) and 4h22m using TV+Breg (71 iterations). Each iteration takes 95s using the optimized CUDA code on a NVIDIA Titan×Maxwell GPU. We computed a maximum of 50 iterations for all iLS+ and TV+ except for TV+Breg. A total of 5 Bregman iterations were performed. Within each Bregman iteration, the TV+ iterations were stopped when either relative or normalised residual falls below a tolerance of 0.01. A more detailed discussion of the stopping condition can be found in [56]. In order to reduce the computation time, deep learning provides opportunities to accelerate the forward model calculations by enabling the use of fast but approximate forward models [57] or increasing the rate of convergence [58].

The current study is expected to inform the practical application of model-based iterative image reconstruction approaches for improving photoacoustic image quality when using finite aperture planar detection geometries.

Acknowledgment

The authors would like to thank Bolin Pan (King's College London) for assistance with the reconstruction script.

References

[1] L. V. Wang, "Multiscale photoacoustic microscopy and computed tomography," *Nature Photon.*, vol. 3, no. 9, pp. 503–509, Sep. 2009, doi: 10.1038/nphoton.2009.157.

[2] P. Beard, "Biomedical photoacoustic imaging," *Interface Focus*, vol. 1, no. 4, pp. 602–631, Aug. 2011, doi: 10.1098/rsfs.2011.0028.

[3] L. Nie and X. Chen, "Structural and functional photoacoustic molecular tomography aided by emerging contrast agents," *Chem. Soc. Rev.*, vol. 43, no. 20, pp. 7132–7170, 2014, doi: 10.1039/c4cs00086b.

[4] J. Xia and L. V. Wang, "Small-animal whole-body photoacoustic tomography: A review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1380–1389, May 2014, doi: 10.1109/TBME.2013.2283507.

[5] J. Yao and L. V. Wang, "Photoacoustic brain imaging: From microscopic to macroscopic scales," *Neurophotonics*, vol. 1, no. 1, May 2014, Art. no. 011003, doi: 10.1117/1.nph.1.1.011003.

[6] A. P. Jathoul et al., "Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter," *Nature Photon.*, vol. 9, no. 4, pp. 239–246, Mar. 2015, doi: 10.1038/nphoton.2015.22.

[7] O. Ogunlade et al., "Monitoring neovascularization and integration of decellularized human scaffolds using photoacoustic imaging," *Photoacoustics*, vol. 13, pp. 76–84, Mar. 2019, doi: 10.1016/j.pacs.2019.01.001.

[8] O. Ogunlade et al., "In vivo three-dimensional photoacoustic imaging of the renal vasculature in preclinical rodent models," *Amer. J. Physiol.-Renal Physiol.*, vol. 314, no. 6, pp. F1145–F1153, Jun. 2018, doi: 10.1152/ajprenal.00337.2017.

[9] S. P. Johnson et al., "Photoacoustic tomography of vascular therapy in a preclinical mouse model of colorectal carcinoma," *Proc. SPIE*, vol. 8943, Mar. 2014, Art. no. 89431R, doi: 10.1117/12.2039843.

[10] L. Lin et al., "High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation," *Nature Commun.*, vol. 12, no. 1, pp. 1–10, Feb. 2021, doi: 10.1038/s41467-021-21232-1.

[11] S. Zackrisson, S. M. W. Y. van de Ven, and S. S. Gambhir, "Light in and sound out: Emerging translational strategies for photoacoustic imaging," *Cancer Res.*, vol. 74, no. 4, pp. 979–1004, Feb. 2014, doi: 10.1158/0008-5472.CAN-13-2387.

[12] A. Taruttis and V. Ntziachristos, "Advances in real-time multispectral optoacoustic imaging and its applications," *Nature Photon.*, vol. 9, no. 4, pp. 219–227, Mar. 2015, doi: 10.1038/nphoton.2015.29.

[13] A. A. Plumb, N. T. Huynh, J. Guggenheim, E. Zhang, and P. Beard, "Rapid volumetric photoacoustic tomographic imaging with a Fabry–Pérot ultrasound sensor depicts peripheral arteries and microvascular vasomotor responses to thermal stimuli," *Eur. Radiol.*, vol. 28, no. 3, pp. 1037–1045, Mar. 2018, doi: 10.1007/s00330-017-5080-9.

[14] P. Hai et al., "Label-free high-throughput photoacoustic tomography of suspected circulating melanoma tumor cells in patients in vivo," *J. Biomed. Opt.*, vol. 25, no. 3, Mar. 2020, Art. no. 036002, doi: 10.1117/1.jbo.25.3.036002.

[15] L. Lin, X. Tong, P. Hu, M. Invernizzi, L. Lai, and L. V. Wang, "Photoacoustic computed tomography of breast cancer in response to neoadjuvant chemotherapy," *Adv. Sci.*, vol. 8, no. 7, Apr. 2021, Art. no. 2003396, doi: 10.1002/advs.202003396.

[16] M. Anastasio, J. A. Z. M. Anastasio, J. Zhang, and M. A. Anastasio, "Image reconstruction in photoacoustic tomography with truncated cylindrical measurement apertures," *Proc. SPIE*, vol. 6086, Mar. 2006, Art. no. 647623, doi: 10.1117/12.647623.

[17] J. Frikel and E. T. Quinto, "Artifacts in incomplete data tomography with applications to photoacoustic tomography and sonar," *SIAM J. Appl. Math.*, vol. 75, no. 2, pp. 703–725, Apr. 2015, doi: 10.1137/140977709.

[18] B. E. Treeby and B. T. Cox, "K-wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.*, vol. 15, no. 2, 2010, Art. no. 021314, doi: 10.1117/1.3360308.

[19] D. Finch and S. K. Patch, "Determining a function from its mean values over a family of spheres," *SIAM J. Math. Anal.*, vol. 35, no. 5, pp. 1213–1240, Jan. 2004, doi: 10.1137/S0036141002417814.

[20] Y. Xu and L. Wang, "Application of time reversal to thermoacoustic tomography," *Proc. SPIE*, vol. 5320, pp. 257–263, Jul. 2004, doi: 10.1117/12.532395.

[21] B. E. Treeby, E. Z. Zhang, and B. T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," *Inverse Problems*, vol. 26, no. 11, Nov. 2010, Art. no. 115003, doi: 10.1088/0266-5611/26/11/115003.

[22] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1097–1110, Jun. 2013, doi: 10.1109/TMI.2013.2254496.

[23] C. Lutzweiler, X. L. Deán-Ben, and D. Razansky, "Expediting model-based optoacoustic reconstructions with tomographic symmetries," *Med. Phys.*, vol. 41, no. 1, Jan. 2014, Art. no. 013302, doi: 10.1118/1.4846055.

[24] K. Wang, R. Su, A. A. Oraevsky, and M. A. Anastasio, "Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography," *Phys. Med. Biol.*, vol. 57, no. 17, pp. 5399–5423, Sep. 2012, doi: 10.1088/0031-9155/57/17/5399.

[25] S. Guan, A. A. Khan, S. Sikdar, and P. V. Chitnis, "Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning," *Sci. Rep.*, vol. 10, no. 1, pp. 1–12, May 2020, doi: 10.1038/s41598-020-65235-2.

[26] R. Shang, R. Archibald, A. Gelb, and G. P. Luke, "Sparsity-based photoacoustic image reconstruction with a linear array transducer and direct measurement of the forward model," *Proc. SPIE*, vol. 24, Dec. 2018, Art. no. 031015, doi: 10.1117/1.JBO.24.3.031015.

[27] M. Ai et al., "Investigation of photoacoustic tomography reconstruction with a limited view from linear array," *J. Biomed. Opt.*, vol. 26, no. 9, Sep. 2021, Art. no. 096009, doi: 10.1117/1.JBO.26.9.096009.

[28] S. Arridge et al., "Accelerated high-resolution photoacoustic tomography via compressed sensing," *Phys. Med. Biol.*, vol. 61, no. 24, pp. 8908–8940, Dec. 2016, doi: 10.1088/1361-6560/61/24/8908.

[29] N. Huynh, E. Zhang, M. Betcke, S. Arridge, P. Beard, and B. Cox, "Single-pixel optical camera for video rate ultrasonic imaging," *Optica*, vol. 3, no. 1, p. 26, Jan. 2016, doi: 10.1364/optica.3.000026.

[30] Y. Hristova, P. Kuchment, and L. Nguyen, "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media," *Inverse Problems*, vol. 24, no. 5, Oct. 2008, Art. no. 055006, doi: 10.1088/0266-5611/24/5/055006.

[31] Y. Xu and L. V. Wang, "Time reversal and its application to tomography with diffracting sources," *Phys. Rev. Lett.*, vol. 92, no. 3, Jan. 2004, Art. no. 033902, doi: 10.1103/PhysRevLett.92.033902.

[32] B. T. Cox, S. Kara, S. R. Arridge, and P. C. Beard, "K-space propagation models for acoustically heterogeneous media: Application to biomedical photoacoustics," *J. Acoust. Soc. Amer.*, vol. 121, no. 6, p. 3453, 2007, doi: 10.1121/1.2717409.

[33] G. Paltauf, J. A. Viator, S. A. Prahl, and S. L. Jacques, "Iterative reconstruction method for three-dimensional optoacoustic imaging," *Proc. SPIE*, vol. 4256, p. 138, Jun. 2001, doi: 10.1117/12.429299.

[34] L. Ding, X. L. Deán-Ben, C. Lutzweiler, D. Razansky, and V. Ntziachristos, "Efficient non-negative constrained model-based inversion in optoacoustic tomography," *Phys. Med. Biol.*, vol. 60, no. 17, pp. 6733–6750, Aug. 2015, doi: 10.1088/0031-9155/60/17/6733.

[35] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992, doi: 10.1016/0167-2789(92)90242-F.

[36] S. R. Arridge, M. M. Betcke, B. T. Cox, F. Lucka, and B. E. Treeby, "On the adjoint operator in photoacoustic tomography," *Inverse Problems*, vol. 32, no. 11, Oct. 2016, Art. no. 115012, doi: 10.1088/0266-5611/32/11/115012.

[37] Y. Dong, T. Görner, and S. Kunis, "An algorithm for total variation regularized photoacoustic imaging," *Adv. Comput. Math.*, vol. 41, no. 2, pp. 423–438, Jun. 2014, doi: 10.1007/S10444-014-9364-1.

[38] J. Meng, L. V. Wang, L. Ying, D. Liang, and L. Song, "Compressed-sensing photoacoustic computed tomography in vivo with partially known support," *Opt. Exp.*, vol. 20, no. 15, pp. 16510–16523, Jul. 2012, doi: 10.1364/OE.20.016510.

[39] M. Burger and S. Osher, "A guide to the TV Zoo," in *Level Set PDE Based Reconstruction Methods Imaging*, M. Burger and S. Osher, Eds. Cham, Switzerland: Springer, 2013, pp. 1–70.

[40] L. Bregman, "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming," *USSR Comput. Math. Math. Phys.*, vol. 7, no. 3, pp. 200–217, Jan. 1967, doi: 10.1016/0041-5553(67)90040-7.

[41] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.

[42] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009, doi: 10.1109/TIP.2009.2028250.

[43] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, May 2011, doi: 10.1007/s10851-010-0251-1.

[44] Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Exp.*, vol. 17, no. 22, Oct. 2009, Art. no. 20178, doi: 10.1364/oe.17.020178.

[45] S. L. Jacques, "Optical properties of biological tissues: A review," *Phys. Med. Biol. Phys. Med. Biol*, vol. 58, pp. 37–61, May 2013, doi: 10.1088/0031-9155/58/11/R37.

[46] E. Zhang, J. Laufer, and P. Beard, "Backward-mode multiwavelength photoacoustic scanner using a planar Fabry–Pérot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues," *Appl. Opt.*, vol. 47, no. 4, pp. 561–577, 2008, doi: 10.1364/AO.47.000561.

[47] M. Lepore and I. Delfino, "Intralipid-based phantoms for the development of new optical diagnostic techniques," *Open Biotechnol. J.*, vol. 13, no. 1, pp. 163–172, Dec. 2019, doi: 10.2174/187407019013016316.

[48] R. Ellwood, F. Lucka, E. Z. Zhang, P. C. Beard, and B. T. Cox, "Photoacoustic imaging with a multi-view Fabry–Pérot scanner," *Proc. SPIE*, vol. 10064, Mar. 2017, Art. no. 100641F, doi: 10.1117/12.2252728.

[49] N. Huynh, O. Ogunlade, E. Zhang, B. Cox, and P. Beard, "Photoacoustic imaging using an 8-beam Fabry–Pérot scanner," *Proc. SPIE*, vol. 9708, Mar. 2016, Art. no. 97082L, doi: 10.1117/12.2214334.

[50] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. Annu. ACM-SIAM Symp. Discret. Algorithms*, vols. 7–9, 2007, pp. 1027–1035.

[51] T. Birsan and D. Tiba, "One hundred years since the introduction of the set distance by Dimitrie Pompeiu," in *Proc. Syst. Modeling Optim., 22nd IFIP TC7 Conf. Held From*, vol. 199, Jul. 2006, pp. 35–39, doi: 10.1007/0-387-33006-2_4.

[52] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inverse Problems*, vol. 19, no. 6, pp. S165–S187, Nov. 2003, doi: 10.1088/0266-5611/19/6/059.

[53] B. Cox, J. G. Laufer, S. R. Arridge, and P. C. Beard, "Quantitative spectroscopic photoacoustic imaging: A review," *J. Biomed. Opt.*, vol. 17, no. 6, 2012, Art. no. 061202, doi: 10.1117/1.JBO.17.6.061202.

[54] D. M. Strong and T. F. Chan, "Spatially and scale adaptive total variation based regularization and anisotropic diiusion in image processing," UCLA Math., CAM Rep. 96-46, 1996.

[55] "Regularization parameter selection methods," in *Computational Methods for Inverse Problems*. Philadelphia, PA, USA: SIAM, 2002, pp. 97–127.

[56] T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a FASTA implementation," 2014, *arXiv:1411.3406*.

[57] A. Hauptmann et al., "Approximate k-space models and deep learning for fast photoacoustic reconstruction," in *Proc. Int. Workshop Mach. Learn. Med. Image Reconstruction* in Lecture Notes in Computer Science, vol. 11074, 2018, pp. 103–111, doi: 10.1007/978-3-030-00129-2_12/FIGURES/6.

[58] A. Hauptmann et al., "Model-based learning for accelerated, limited-view 3-D photoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1382–1393, Jun. 2018, doi: 10.1109/TMI.2018.2820382.